基于群智感知的街景变化检测方法 *

钟委钊 a，陈荟慧 b†
(佛山科学技术学院 a. 机电工程与自动化学院; b. 电子信息工程学院，广东 佛山 528225)

摘 要：移动群智感知数据包含的图像和时空情境信息可用于检测街景图像变化，但是群智感知数据通常是低质和不规范的。为了准确检测街景发生的变化，主要解决由拍摄视角差异引起的数据低质问题。首先，针对大视差问题采用图像配准方法初步对齐图像并提取出配准特征点；然后，基于配准特征点分布从图像中提取感兴趣区域；第三，针对差值图像的误检内容，提出基于面积和多特征点的筛选法去除误检区域；最后，结合边缘检测和超像素分割算法提取完整的变化对象。将所提方法与 MDFNet 方法进行比较，实验结果显示：当街景发生变化时，所提方法的 F1-Measure 值为 55.8%，增长 6%，错误率为 10.8%，降低 24%；当街景无变化时，所提方法的错误率为 2.8%，下降 28%。

关键词：移动群智感知；图像配准；变化检测；街景；电子地图
中图分类号：TP381　　doi: 10.19734/j.issn.1001-3695.2022.02.0082

Change detection method of street view image based on crowdsensing

Zhong Weizhao[a],Chen Huihui[b†]
(*a. Dept. of Electromechanical Engineering & Automation*, *b. Dept. of Electronics & Information Engineering*, *FoShan University*, *FoShan Guangdong* 528225, *China*)

Abstract: The images and spatiotemporal contextual information contained in mobile crowdsensing data can be used to detect changes in street view images, but crowdsensing data are usually low-quality and non-standard. In order to accurately detect the changes in the street view, this paper mainly solves the problem caused by low-quality datum which have been collected from different shooting perspectives. First, for the problem of large parallax, the image registration method is used to initially align the images and extract the registration feature points; then, the region of interest is extracted from the image based on the distribution of the registration feature points; third, aiming at the error detection content of the difference image, a screening method based on area and multi-feature points is proposed to remove the error detection region; finally, the complete changing object is extracted by combining edge detection and superpixel segmentation algorithm. Compared with MDFNet method, when the image pairs change, the F1-measure value and error rate of this method are 55.8% and 10.8%, and they have increased by 6% and decreased by 24% respectively. When there are no changes, the error rate of this method is only 2.8% and it is reduced by about 28%.

Key words: mobile crowdsensing;image registration;change detection;street view;electronic map

0 引言

随着我国城乡建设的发展，道路建设日新月异，人们对于出行辅助工具软件的需求越来越丰富，各类电子地图给人们的出行带来便捷。目前人们的出行已经基本依赖于电子地图。在电子地图的基础上，谷歌街景、百度街景等利用虚拟现实技术将街道全景照片映射到电子地图，让人们足不出户便可以通过电子地图和街景游遍全球。

街景不但为人们的出行提供了便利和安全保障，比如路面缺陷预警[1,2]，也为城市化建设提供了展示平台[3]。专业街景车是街景图像采集和更新的主要工具，但定时更新街景的工作量非常大。移动群智感知使用人们携带的移动智能设备(例如，智能手机)来感知和收集信息，数据采集成本低且覆盖度广，可用于街景图像更新[4]、路况监测[5,6]、环境质量检测[7]等。基于移动群智感知的街景变化检测具备很好的时空优势，图像采集成本低、更新快且范围广[8,9]，能够实现街景地图的短周期更新，但是，群智感知数据存在低质、冗余、差异大、不规范等缺点[8]。

街景变化情况通常通过对比拍于不同时间的街景照片的差异进行检测。基于群智感知手段收集街景数据比街景车更灵活，但是，照片受角度差别、光线差别等因素影响，导致数据集质量比较低，因此，根据图像差异判断街景图像的变化比较困难。首先，由于采集人位置不固定，两张不同的街景图像可能是由于视角差异引起的图像差异，从而导致判断失效。其次，图像的差异具体表现在像素上，而街景变化以物体(对象)为单位出现或者消失，需要从像素级别的差异信息中扩展并获取具体的变化对象。

为了解决以上两个问题，本文的贡献包括：a)提出端到端的针对大视差街景图像对变化对象检测的算法框架，通过引入图像配准方法解决视差问题，并提出基于多特征点空间分布信息的消除错误检测信息的算法；b)根据物体边界的连续性，结合边缘检测和超像素分割算法提取具体的变化对象；c)以校园为测试场景，采集并构建了校园街景图像数据集，通过对比最新的相关工作，证明方法的有效性。

1 相关工作

街景变化的检测方法已有大量的研究[10~24]。根据街景变化类型、方法和用途，可以将它们分类为二维图像变化检测

收稿日期：2022-02-21；**修回日期：**2022-04-19　**基金项目：**国家自然科学基金资助项目(61972092，61902068)

作者简介：钟委钊(1997-)，男，广东茂名人，硕士研究生，主要研究方向为智能感知、图像处理；陈荟慧(1978-)，女(通信作者)，河南洛阳人，教授，硕导，博士，主要研究方向为智能感知、城市大数据(hhchen@fosu.edu.cn).

和三维场景建模检测。

二维图像变化检测是主要的街景变化检测方法[10~19]。街景变化的传统检测方法是根据在不同时间捕获的一组图像创建街景的外观模型，并将其与新捕获的查询图像进行比较以检测变化。这类研究的主要关注点是如何处理不相关的外观变化，如光照差异[10]。为了能从不相关的外观变化中更准确地检测出真实的变化，Zhao 等[11]提出了一种孪生编码器-解码器的网络结构，实现语义级别的变化检测；Chen 等[12] 在编码器-解码器的网络结构上添加了动态感知时间注意模块，并结合水平和垂直的并发注意模块对检测结果进行精细化；也有研究提出从不同的网络层中提取特征进行融合，检测从整体到局部的变化区域[13]。

在视角差异的问题上，也有大量的研究。Sakurada 等[14] 提出了一种利用卷积神经网络特征与超像素分割相结合的变化检测方法。其他基于神经网络的检测方法还有：结合卷积网络和反卷积网络的多层特征融合实现不同尺度变化信息获取的 ChangeNet[15]；结合语义分割模型和图割的街景影像变化检测方法[16]；通过融合多重差异特征以获得不同尺度变化信息的 MDFNet(Multiple Difference Features Network)[17] 方法；使用光流估计法配准图像对得到掩码图像，再结合原图像结构信息得到变化区域的 Mask-CDNet(Mask based Pixel Change Detection Network)[18]。

除了视角差异外，为了处理放缩、旋转引起的图像差异，Guo 等提出了全卷积孪生度量网络 CosimNet(fully Convolutional Siamese metric Network)，并加入了阈值对比损失惩罚噪声变化[19]。

在三维领域中的也有很多解决方法[20~22]。它们采集多组街景图像构建多视图立体(Multi View Stereo，MVS)。通过对比新模型和原始模型之间的差异来检测 MVS 的变化，并局部更新 MVS。这种比较的准确性取决于可获得的 MVS 重建的质量。场景的三维模型通常通过使用三维传感器而不是普通相机来创建。有的研究将多传感器融合的即时定位与地图构建(SLAM，Simultaneous Localization And Mapping)和快速密集三维重建流水线连接起来，将粗配准的图像对提供给 DN(Deconvolutional Networks)，用于像素方向的变化检测[20]。有的研究通过 3D 模型将图像重新投影到另一个图像上来发现图像对之间的变化，并使用多个图像组合解决检测过程中产生的歧义，以便估计变化的 3D 位置[21]。这些方法是为维护/更新现有的三维城市模型而设计的。

还有另一种类型的研究，使用大量街景的多视图图像并通过运动结构的方法创建其时空模型。有研究利用几十年来拍摄的大量城市图像，实现特定类型的时间推断，例如估计建筑的建造时间[23,24]。在思想上，他们的方法都是使用互联网图像数据集来检测场景的变化，如广告和建筑墙上绘画的变化。假设有足够数量的场景多视图图像可用，这方法可以使用 SFM(Structure From Motion)来重建场景的三维模型。

基于群智感知的城市街景图像更新主要通过手机拍摄的照片差异识别街景的变化，也是通过比较由参与者在不同时间拍摄的街景图像来检测街景变化。本文提出的方法属于二维图像变化检测方法，区别于传统的检测方法和现有的神经网络方法，本文方法在不进行大量的数据训练和语义分割的情况下，通过多个图像处理模块，检测视差街景图像对的变化内容，并从中提取具体的变化对象。

2　街景变化检测算法

2.1　算法框架

街景变化的检测主要通过对拍照情境(包括位置和拍照角度)相似的两张照片进行差异检测，分为 4 个步骤，包括图像预处理、变化区域提取、变化区域筛选和变化区域扩展，流程图如图 1 所示。

a)图像预处理。包括使用尺寸重构把图像对转换为同一尺寸；使用直方图规定化[25]使图像对的灰度直方图一致；使用图像配准算法对齐街景图像对，本文以基于 SIFT(Scale-Invariant Feature Transform)特征点的 APAP 算法[26]为例；提取涵盖所有配准特征点的最小矩形区域作为感兴趣区域(ROI，Region of Interest)进行研究，减少不必要的计算量。

b)在 ROI 中提取变化区域。包括使用图像减法得到二值差异图；使用基于阈值的多种子生长法提取天空区域，解决天空区域多变的问题[27]；使用孔洞填充法解决配准造成的像素缺失问题；使用形态学操作把变化内容分成多个子区域提取出来。

c)对变化区域进行筛选。包括使用面积筛选法滤除变化内容像素的面积过小的变化子区域；使用多特征点筛选法滤除较多特征点分布的变化子区域。

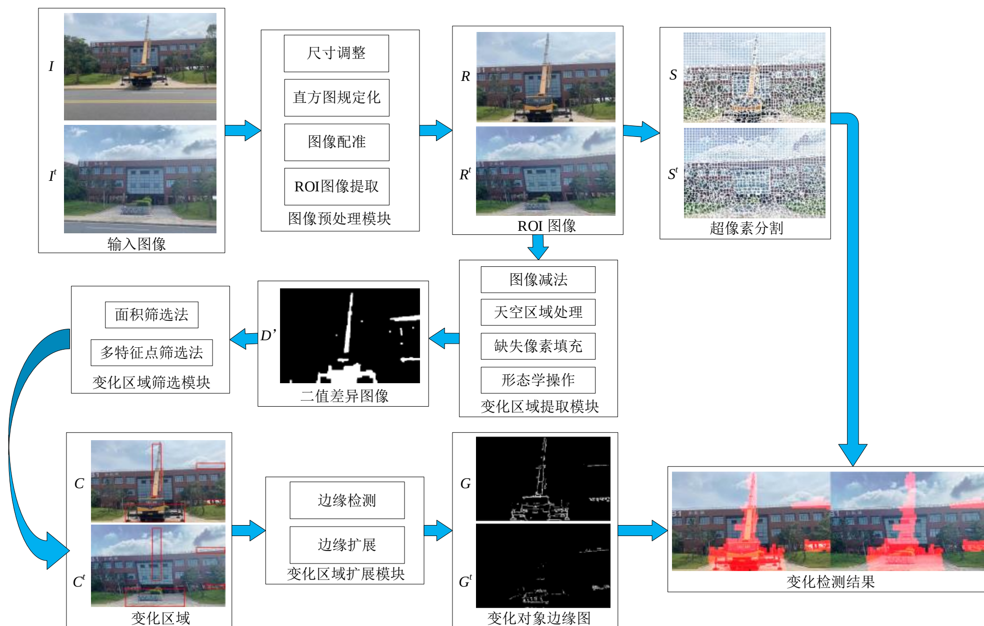


图 1　变化检测算法框架图

Fig. 1　Framework of the proposed change detection method

最后，扩展变化区域，并提取完整的变化对象。包括使用边缘检测法提取变化区域中物体的轮廓；使用边缘扩展法对轮廓进行 4 邻域延伸，得到完整的物体轮廓；最后结合超像素分割法[28]提取出完整的变化对象。

2.2 ROI 图像提取

ROI 图像提取目的在于从图像中分割出用于识别街景变化的图像区域，并剔除其余区域。因为没有特征点分布的区域是不重合的或者没有经过配准的，对这种区域操作会产生大量不必要且错误的运算。

记原街景图像为 I^t，新提交图像为 I。配准街景图像 I 和 I^t 后，得到图像 I^t 映射变换后的 I'。由于 APAP 算法[26]是一种单向配准算法，目标图像 I 在配准过程中不发生形变，所以使用图像 I 中涵盖所有配准特征点的最小矩形区域来提取图像 I' 和 I 的 ROI，记作 R^t 和 R。R^t 和 R 分别是 I' 和 I 的子图像。配准特征点指的是在 APAP 算法中用于估计单应矩阵 H 的 SIFT 特征点。

2.3 街景变化区域提取

街景变化区域的提取过程分为四部分：图像求差、天空消除、修补边缘和区域连通。

经过配准后，图像 R^t 和 R 在相同位置的像素信息可视为是对应的，求得的差值图像即为变化图像。为了简化计算，使用式(1)把彩色图像 R^t 和 R 灰度化，再对灰度化后的两幅图像求差，得到差值图像 D。使用大津法[29]找到最佳阈值将图像 D 转换为二值图像 D'，像素值为 1 的区域是两幅图像的差别区域，值为 0 的区域是相同区域。

$$gy = 0.299\times r + 0.578\times g + 0.114\times b \tag{1}$$

其中，r, g, b 分别表示红、绿、蓝三通道值，gy 表示灰度值。

通常，两张照片采集的时间是不同的，不同的天气和云彩可能被误检为差别区域，因此，消除天空区域也是街景差别检测的重要工作。基于阈值分割的种子生长法是提取天空区域的常用方法，但是使用单一种子点容易导致分割出来的天空区域不完整，所以本文引用多种子生长法来提取天空区域[21]。

从图像 R^t 和 R 中提取出天空区域的二值化掩膜图像分别记为 M^t 和 M，其中，像素值为 1 的区域是天空区域，像素值为 0 的区域是非天空区域。对图像 M^t 和 M 使用孔洞填充法后，使用式(2)更新差异图 D' 的天空部分的像素。

$$D'_{x,y} \leftarrow \left|M^t_{x,y} - M_{x,y}\right|, if\ M^t_{x,y} \vee M_{x,y} = 1 \tag{2}$$

其中，$M_{x,y}$ 表示图像 M 在(x, y)处的像素值。网格配准的图像边缘区域可能由于大形变后边界像素缺失导致出现 1 值像素块，而这种黑色像素块在图像相减后可能被当作变化区域提取出来，所以需要在差异图 D' 中提取变化内容前填充像素缺失的区域。填充方法如式(3)所示。

$$D'_{x,y} \leftarrow 0, if\ (D'_{x,y} = 1) \wedge (R^t_{x,y} \vee R_{x,y} = 0) \tag{3}$$

其中，$D_{x,y}$ 表示 D 在(x, y)处的像素值。

在二值差异图像 D' 中，可能出现变化对象的轮廓并不清晰，内部存在空洞区域，相邻对象存在牵连的问题，因此，存在黑点噪声和白点噪声。在分别使用开运算和闭运算后，消除图像中的黑白噪声，得到多个无噪声的连通区域。最后，通过二值差异图像像素的 4 邻域得到连通区域，此连通区域即为初始的变化区域，连通区域的个数为变化区域的个数，这些变化区域集合记为 C_1。

2.4 变化区域筛选

C_1 中存在部分由于图像配准不精确造成的误检变化区域，需要消除这些错误检测的区域。

误检变化区域类型主要包括：a)由于微小形变带来的小面积误检变化区域；b)由视角差导致遮掩问题所带来的误检变化区域。针对类型 1，使用变化区域面积过滤法；针对类型 2，使用基于多特征点的网格筛选法。流程图如图 2 所示。

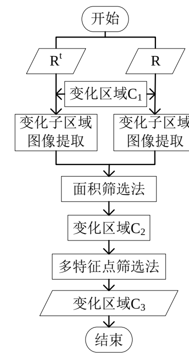


图 2　变化区域提取流程图

Fig. 2　Flowchart of change extraction

2.4.1 面积筛选法

本文关注的是街景图像上的发生较大变化区域的情况，未严格对齐的差异像素点组成的细小面积区域是可以忽略的。面积过小的变化区域可以滤除，过滤阈值记作 th(本文 th=0.03)，使用式(4)把小面积候选区域滤除后，得到新的变化区域集合 C_2。

$$C_2 \leftarrow C_1 - c, if\ \frac{S(c)}{S(R)} < th, \forall c \in C_1 \tag{4}$$

其中，$S(k)$表示变化区域 k 的面积。

2.4.2 多特征点筛选法

变化区域内的图像匹配特征点数量较少，因此，匹配特征点数量可以用来判断变化区域是否为真。但是，变化区域内的物体缺少匹配特征点，仅使用 SIFT 特征点在数量上是不足够的。并且由于视角差异的存在，在特征描述上要求特征描述子需具有尺寸不变性，因此，本文使用 SIFT 和 ORB (Oriented FAST and Rotated BRIEF)相结合的多特征点。多特征点方法在特征点数量增多的同时，能增加特征点的多样性。为了保证匹配特征点的质量，使用 RANSAC 算法滤除错误匹配的特征点对。以下特征点提取默认使用 RANSAC 算法。

由于匹配特征点在图像中的空间分布是不均匀的。例如，如果大量匹配特征点聚集，那么一个包含这些聚集特征点的较大区域内的匹配特征点数量仍然非常多。所以，直接使用匹配特征点数量容易造成误判，因此，本文采用基于网格划分的特征点数量比对方法。

由于多特征点匹配过程中不可避免受到视角差的影响，因此本文提出分别提取图像 R^t 和 R、I^t 和 I 的 SIFT 特征和 ORB 特征并组成新的多特征集合，在保证特征点数量的同时，增大了特征匹配的“视野”。

变化区域集 C_2 中可能包含了未真实发生变化的区域。针对 C_2 中的任一区域，提取该区域在图像 R^t 和 R 上的子图像，记为 I_1 和 I_2。将子图像划分为 5×5 的网格，用能够覆盖到匹配特征点的网格的数量的倒数表示图像的相似指数 d，如式(5)所示。

$$d_{I_1} = \frac{1}{gn_1}, d_{I_2} = \frac{1}{gn_2} \tag{5}$$

其中，gn_i 表示能够覆盖图像 I_i 的匹配特征点的网格数量。一副图像的相似指数越大，说明它与另一幅图像越相似。

$$thr = \max\left(\frac{1000}{S(c)}, \frac{1}{12}\right), \forall c \in C_2 \tag{6}$$

其中，$S(k)$表示子图像 k 的面积。如果相似指数小于阈值 thr，从 C_2 移除该区域，最后得到过滤后的变化区域集 C_3。

2.5 变化区域扩展

变化区域筛选后，需要通过区域连通和扩展方法获取完整的变化区域，才能提取到发生变化的物体。同一物体的边缘像素点通常具有连续性和相似性，因此，从边缘对图像进

行连通区域扩展，先检测边缘，提取到图像中发生变化的物体的区域后，采用超像素法分割图像。

a)使用 Sobel 算子对 C_3 变化区域的子图像进行边缘检测，得到梯度值图。去除梯度值过小的点后(阈值为 80)，梯度值不等于 0 的像素点作为初始边缘点，放入集合 G 中。然后，在原图像中分别遍历 G 中的点，将 G 中各点的连通点分别逐步放入 G 中。点 A 是点 B 的连通点需满足三个条件：点 A 与点 B 相邻；在梯度值图中，点 A 处与点 B 处的梯度值之差小于阈值 thg(本文 thg=50)；以 A 为中心的 21×21 个像素内没有匹配特征点。最后，G 中的点组成的区域即为连通区域。

b)基于超像素分割法[22]获取完整的变化图像内容。首先，随机选取 ns 个种子点，使用超像素法把原图像分割成 ns 个形状不规则的像素块。然后，针对集合 G 中的每一个像素点 g_i，使用式(7)计算 g_i 它半径为 q(本文 q=10)像素的范围内的种子点的距离，找到距离其最近的种子点 s_j，将点 g_i 和点 s_j 所在像素块合并。最后以合并后的像素块的中心点作为新的种子点，进行下一轮迭代。

$$s_j = \arg\min\left(\left|b_{g_i} - b_{s_j}\right| + \left|l_{g_i} - l_{s_j}\right| \times k\right), \left|l_{g_i} - l_{s_j}\right| \le q \tag{7}$$

其中，b_p 表示 p 的 RGB 颜色向量，l_p 表示像素点 p 的位置向量，k 为权重系数。

迭代 n 次后，集合 G 中每个像素点对应的像素块即为最终标注的街景变化区域。

3　实验结果

3.1　数据集

实验以校园为街景变化检测对象，招募 5 名志愿者按照拍摄地点和拍摄姿态采集数据。拍摄姿态包括方位角和俯仰角。通过利用手机传感器，数据采样 App 能够基本保证拍摄过程中方位角和俯仰角的偏差角度不大于 3，GPS 定位距离不大于 1 米。图像采集时间跨度不超过 3 个月。实验总共采集了 100 组校园街景图像，包含了 15 个场景，用于实验的照片尺寸统一压缩为 640×480。在实验中，100 组街景图像被分为两组。第一组是街景发生变化的图像，共 60 对照片，记为 G1；第二组为街景未发生变化的图像，共 40 对照片，记为 G2。为评估和对比方法的效果，实验还对数据集进行了标注，0(黑色区域)表示非变化区域，而变化区域则保留原有的颜色信息。检测数据主要采集了路面和楼面物品的改变，例如楼面装饰、障碍物、汽车的出现/消失。

3.2　实验参数

经过多次实验，本文选择了实验结果最好一组参数作为文本的实验参数。部分参数已在文中直接给出，在这里不再阐述。在街景变化区域提取中，开运算的卷积核大小为 5×5，闭运算的卷积核大小为 15×15。在变化区域筛选中，SIFT 特征点匹配参数与文献[26]中一致；ORB 相似特征点匹配的距离阈值设为 0.9；APAP 算法匹配的比率阈值设为 0.8，即最优匹配特征向量距离和次优匹配特征向量距离的比值小于等于 0.8；除此之外，RANSAC 中数量最大值的可信度设为 90，投影位置和相应位置之间的最大像素距离设为 60。在变化区域扩展中，超像素分割法的初始种子点数 ns 为 3000，权重系数 k 为 1.5，种子点的迭代次数 n 为 10。

3.3　评价指标和基准方法

实验评价采用了 4 个评价指标，分别是精确率、召回率、F1-measure 和错误率。

精确率指从 ROI 图像(有效检测区域，即两幅输入图像的重叠区域)中实验检测到的正确变化区域面积占检测到的变化区域总面积的比例；召回率指从 ROI 图像中检测到的正确变化区域面积占真实变化区域面积的比例；F1-Measure 值根据精确率和召回率计算；错误率指从 ROI 图像中提取的错误变化区域面积占真实的非变化区域面积的比例。

为了验证所提算法的有效性，使用大津法[29]结合图像差值法作为基准方法，并与工作相近的文献[15~17]中所提方法进行比较。

3.4　实验结果分析

本文在训练文献[15]的 ChangeNet 模型和文献[17]的 MDFNet 模型时，除了使用文献[17]中提到的 PCD 数据集，还添加了 70 对的本文采集的图像，其中 G1 组图像 42 对，G2 组图像 28 对，并对图像进行了数据增广。而文献[18]的 DeeplabV3+网络使用 Camvid 数据集进行训练。以下的可视化结果从 50 epoch 的训练中取测试集损失最小的模型所得。

本文方法在数据集 G1 和 G2 下的平均精确率、召回率、F1-measure 值和错误率的实验结果统计如表 1 所示。数据集 G2 中的图像对未发生变化，精确率、召回率和 F1-Measure 无效，不予评价。

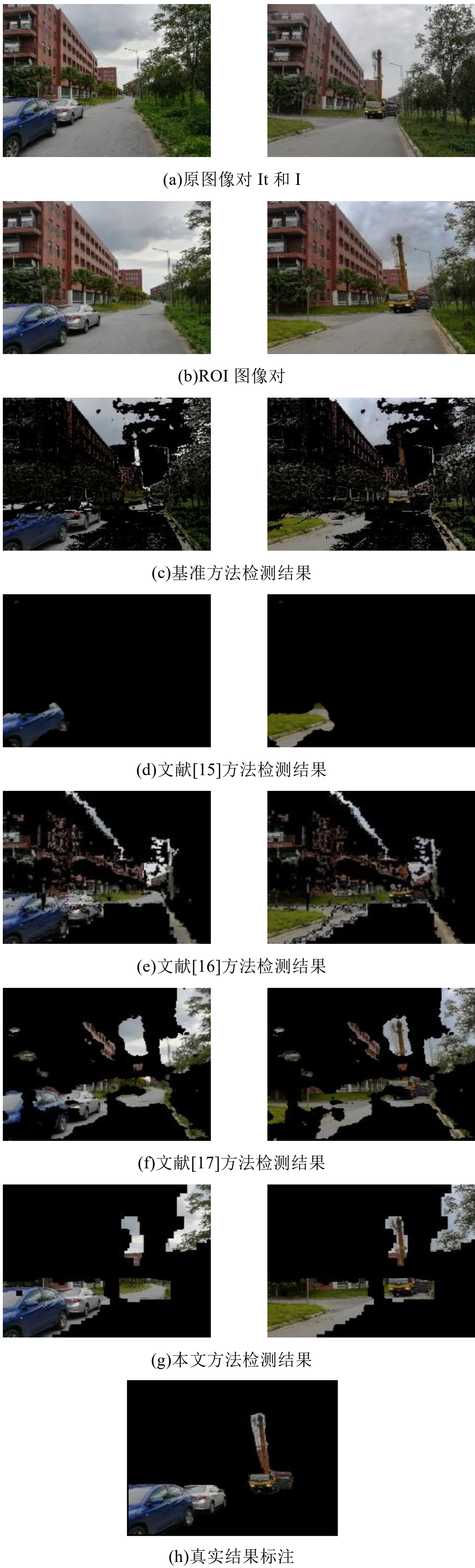
表 1　实验结果

Tab. 1　Experimental results

实验方法	精确率	召回率	F1-measure	错误率
基准方法(G1)	50.6	11.9	17.8	45.6
文献[15]方法(G1)	81.9	22.8	31.6	1.7
文献[16]方法(G1)	72.2	35.7	48.5	30.7
文献[17]方法(G1)	71.4	38.3	49.9	34.6
本文方法(G1)	**75.1**	**47.6**	**55.8**	**10.8**
基准方法(G2)	—	—	—	48.6
文献[15]方法(G2)	—	—	—	0.9
文献[16]方法(G2)	—	—	—	31.8
文献[17]方法(G2)	—	—	—	30.5
本文方法(G2)	—	—	—	**2.8**

结果显示，G1 组数据采用本文方法得到了 55.8%的 F1-measure 值，比基准方法、文献[15]方法、文献[16]方法和文献[17]方法得到的分别增长了约 38%、24%、7%和 6%；且本文方法获得的平均错误率为 10.8%，较基准方法、文献[16]方法和文献[17]方法分别下降了约 35%、20%和 24%，而较文献[15]方法增长了约 9%。G2 组数据采用本文方法得到了 2.8%的平均错误率，比基准方法、文献[16]方法和文献[17]方法得到的分别下降了约 46%、29%和 28%，而较文献[15]方法增长了约 2%。为了更直接地比较，图 3~5 是在数据统计过程中随机选取未参与训练的三个图像对作为案例进行可视化比较。当图像对中存在变化内容时，如图 3 和 4 所示，分别把方法检测结果(c)(d)(e)(f)(g)和真实结果标注(h)进行比较，可知本文方法较对比方法能更精确地检测出两幅街景图像中的变化对象。当图像对中不存在变化内容时，如图 5 所示，本文方法无错检内容，除文献[15]的其他对比方法的检测结果仍存在大片错检区域。

结合指标和可视化结果可得，本文方法整体上优于对比方法，且能更精确地提取变化对象。文献[15]方法在两个数据集上获得了最低的错误率和较低的召回率，对局部的变化检测上比较准确，但对较大的变化对象的提取能力较弱。文献[16]的方法虽然通过语义分割对视差具有一定的鲁棒性，但对于图像配准后产生的扭曲，其逐像素比较得到变化结果的方法无法得到较好的检测结果，而且错误率会更高，在复杂的街景环境下，语义分割的效果不完全正确也是原因之一。文献[17]的方法效果仅次于本文的方法，其提取的多层不同尺寸的差异特征能较好地处理视差问题，但是细节部分的处理能力欠佳，且误检率较高。而本文借助配准特征点对的分布信息，能较好地处理由于像素未能精确对齐而产生的差异问题。最后基于物体边缘的连续性这一特性，以提取出具体的变化对象为导向，提高精确率、召回率和 F1-measure，并减少错误率。


(a)原图像对 It 和 I

(b)ROI 图像对

(c)基准方法检测结果

(d)文献[15]方法检测结果

(e)文献[16]方法检测结果

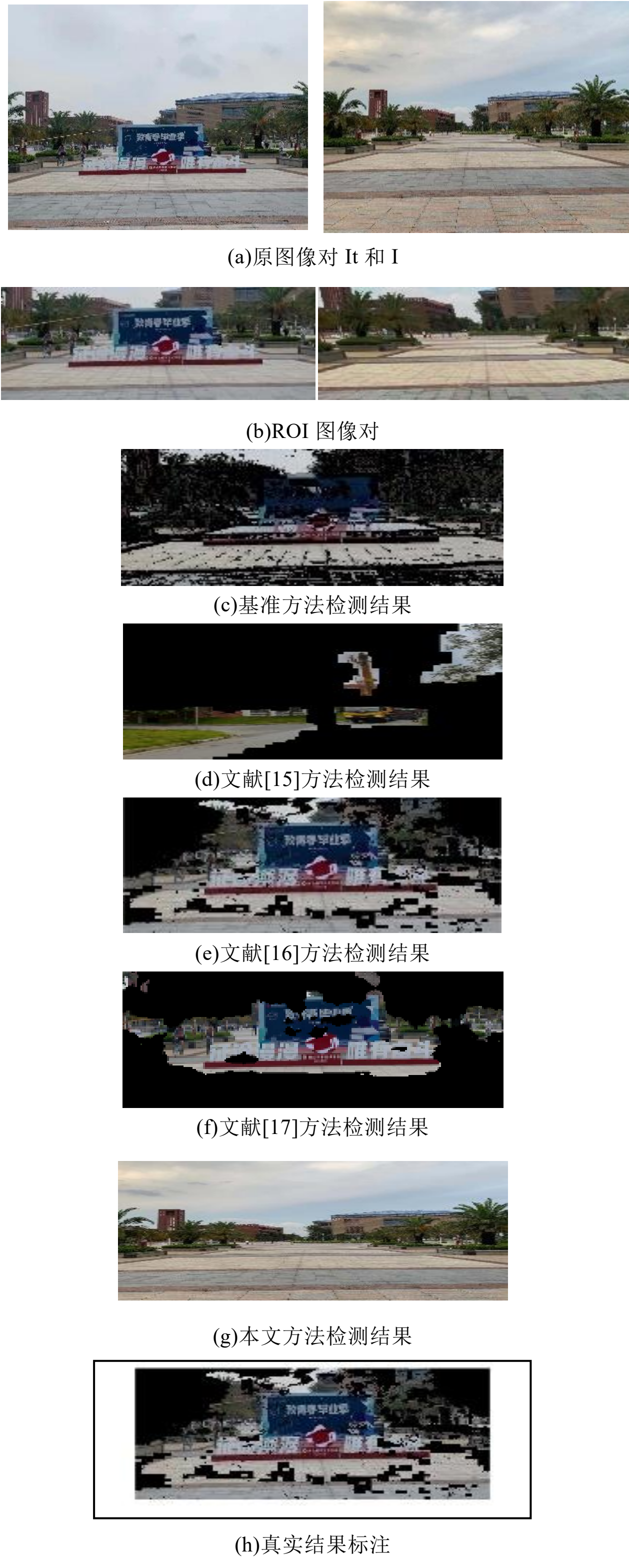
(f)文献[17]方法检测结果

(g)本文方法检测结果

(h)真实结果标注

图 3　G1 组实验案例 A：汽车的出现/消失

Fig. 3　Experimental case A in G1: appearance/disappearance of cars


(a)原图像对 It 和 I

(b)ROI 图像对

(c)基准方法检测结果

(d)文献[15]方法检测结果

(e)文献[16]方法检测结果

(f)文献[17]方法检测结果

(g)本文方法检测结果

(h)真实结果标注

图 4　G1 组实验案例 B：障碍物的出现/消失

Fig. 4　Experimental case B in G1: appearance/disappearance of obstacles

4　结束语

本文提出了一种基于群智感知数据检测街景变化的框架和方法。实验采用由志愿者拍摄的校园街景图像数据集，实验指标值和实验结果显示该方法目前能够较准确地检测出街景变化的内容。本文的方法比较依赖于图像配准特征点对的分布信息，通过多特征点组合的方法来提高配准精度，但是增加了计算消耗，本文使用的 APAP 图像配准算法还有待改进。未来，本文将研究如何结合特征点的空间分布信息和轻量级的神经网络，更好地完成基于群智感知的街景图像变化检测任务。




(a)原图像对 It 和 I



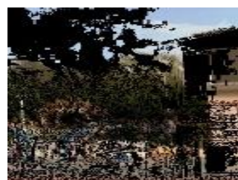

(b)ROI 图像对



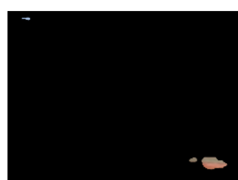
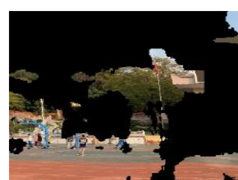

(c) 基准方法检测结果



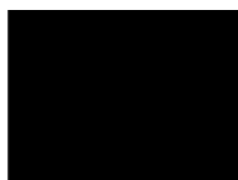

(d)文献[15]方法检测结果



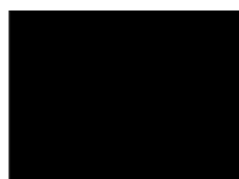

(e)文献[16]方法检测结果

(f)文献[17]方法检测结果(存在变化)




(g)本文方法检测结果(无变化)



(h)真实结果标注(无变化)

图 5　G2 组实验案例 C：无明显变化内容

Fig. 5　Experimental case C in G2: no expected change

参考文献：

[1] Huang Jiong, Hu Sheng, Wang Yun, *et al.* GrabView: a scalable street view system for images taken from different devices [C]// Proc of the 5th International Conference on Multimedia Big Data (BigMM) . IEEE, 2019: 302-304.

[2] Vishnani V, Adhya A, Bajpai C, *et al.* Manhole detection using image processing on Google Street View imagery [C]// Proc of the 3rd International Conference on Smart Systems and Inventive Technology (ICSSIT) . IEEE, 2020: 684-688.

[3] Xu Chi, Chen Qiao, Liu Jiangchuan, *et al.* Smartphone-based crowdsourcing for panoramic virtual tour construction [C]// Proc of IEEE International Conference on Multimedia & Expo Workshops (ICMEW) . IEEE, 2018: 1-4.

[4] Peng Zhe, Gao Shang, Xiao Bin, *et al.* CrowdGIS: Updating digital maps via mobile crowdsensing [J]. IEEE Transac-tions on Automation Science and Engineering, 2017, 15 (1): 369-380.

[5] Wang Qianru, Guo Bin, Wang Leye, *et al.* Crowdwatch: dynamic sidewalk obstacle detection using mobile crowd sensing [J]. IEEE Internet of Things Journal, 2017, 4 (6): 2159-2171.

[6] Kong Yingying, Yu Zhiwen, Chen Huihui, *et al.* Detecting type and size of road crack with the smartphone [C]// Proc of IEEE International Conference on Computational Science and Engineering (CSE) and IEEE International Conference on Embedded and Ubiquitous Computing (EUC) . IEEE, 2017, 1: 572-579.

[7] Leonardi C, Cappellotto A, Caraviello M, *et al.* SecondNose: an air quality mobile crowdsensing system [C]// Proc of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational. 2014: 1051-1054.

[8] Chen Huihui, Guo Bin, Yu Zhiwen. Measures to improve outdoor crowdsourcing photo collection on smart phones [C]// Proc of IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Compu-ting, Internet of People and Smart City Innovation (Smart-World/SCALCOM/UIC/ATC/CBDCom/IOP/SCI) . IEEE, 2019: 907-915.

[9] Chen Huihui, Cao Yangjie, Guo Bin, *et al.* LuckyPhoto: multi-facet photographing with mobile crowdsensing [C]// Proc of ACM International Joint Conference and Interna-tional Symposium on Pervasive and Ubiquitous Computing and Wearable Computers. 2018: 29-32.

[10] Radke R J, Andra S, Al-Kofahi O, *et al.* Image change detection algorithms: a systematic survey [J]. IEEE Trans on Image Processing, 2005, 14 (3): 294-307.

[11] Zhao Xinwei, Li Haichang, Wang Rui, *et al.* Street-view change detection via siamese encoder-decoder structured convolutional neural networks [C]// VISIGRAPP (5: VISAPP) . 2019: 525-532.

[12] Chen Shuo, Yang Kailun, Stiefelhagen R. DR-TANet: dynamic receptive temporal attention network for street scene change detection [C]// Proc of IEEE Intelligent Vehi-cles Symposium (IV) . IEEE, 2021: 502-509.

[13] Lei Yinjie, Peng Duo, Zhang Pingping, *et al.* Hierarchical paired channel fusion network for street scene change detection [J]. IEEE Trans on Image Processing, 2020, 30: 55-67.

[14] Sakurada K, Okatani T. Change detection from a street image pair using CNN features and superpixel segmenta-tion [C]// BMVC. 2015, 61: 1-12.

[15] Varghese A, Gubbi J, Ramaswamy A, *et al.* ChangeNet: a deep learning architecture for visual change detection [C]// Proc of European

Conference on Computer Vision (ECCV) Workshops. 2018: 0-0.

[16] 李文国, 黄亮, 左小清, 等. 一种结合语义分割模型和图割的街景影像变化检测方法 [J]. 全球定位系统, 2021, 46 (01): 98-104. (Li Wenguo, Huang Liang, Zuo Xiaoqing, *et al.* A street view image change detection method combining semantic segmentation model and graph cuts [J]. GNSS World of China, 2021, 46 (01): 98-104.)

[17] 詹瑞, 雷印杰, 陈训敏, 等. 基于多重差异特征网络的街景变化检测 [J]. 计算机科学, 2021, 48 (02): 142-147. (Zhan Rui, Lei Yinjie, Chen Xunmin, *et al.* Street-view change detection based on multiple difference feature network [J]. Computer Science, 2021, 48 (02): 142-147.)

[18] Bu Shuhui, Li Qing, Han Pengcheng, *et al.* Mask-CDNet: a mask based pixel change detection network [J]. Neuro-computing, 2020, 378: 166-178.

[19] Guo Enqiang, Fu Xinsha, Zhu Jiawei, *et al.* Learning to measure change: fully convolutional siamese metric networks for scene change detection [J]. ArXiv Preprint ArXiv, 2018, 1810. 09111.

[20] Alcantarilla P F, Stent S, Ros G, *et al.* Street-view change detection with deconvolutional networks [J]. Autono-mous Robots, 2018, 42 (7): 1301-1322.

[21] Ku Tao, Galanakis S, Boom B, *et al.* SHREC 2021: 3D point cloud change detection for street scenes [J]. Computers & Graphics, 2021, 99: 192-200.

[22] Taneja A, Ballan L, Pollefeys M. City-scale change detection in cadastral 3D models using images [C]// Proc of IEEE Conference on computer Vision and Pattern Recognition, 2013: 113-120.

[23] Schindler G, Dellaert F. Probabilistic temporal inference on reconstructed 3D scenes [C]// Proc of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2010: 1410-1417.

[24] Matzen K, Snavely N. Scene chronology [C]// Proc of European Conference on Computer Vision. Springer, Cham, 2014: 615-630.

[25] Coltuc D, Bolon P, Chassery J M. Exact histogram specifi-cation [J]. IEEE Trans on Image Processing, 2006, 15 (5): 1143-1152.

[26] Zaragoza J, Chin T J, Brown M S, *et al.* As-projective-as-possible image stitching with moving DLT [C]// Proc of IEEE Conference on Computer Vision and Pattern Recognition. 2013: 2339-2346.

[27] 张燕丽, 柯旭. 结合天空区域检测的图像去雾算法研究与实现 [J]. 重庆工商大学学报 (自然科学版) , 2017, 34 (05): 37-42. (Zhang Yanli, Ke Xu. Research and implementation of image defogging algorithm combined with sky area detection [J]. Journal of Chongqing Technology and Business University (Natural Science Edition) , 2017, 34 (05): 37-42.)

[28] Stutz D, Hermans A, Leibe B. Superpixels: An evaluation of the state-of-the-art [J]. Computer Vision and Image Understanding, 2018, 166: 1-27.

[29] Otsu N. A Threshold selection method from gray-level histograms [J]. IEEE Trans on Systems, Man, and Cybernetics, 1979, 9 (1): 62-66.